

J.-B. Sirven · B. Bousquet · L. Canioni · L. Sarger ·  
S. Tellier · M. Potin-Gautier · I. Le Hecho

## Qualitative and quantitative investigation of chromium-polluted soils by laser-induced breakdown spectroscopy combined with neural networks analysis

Received: 28 October 2005 / Revised: 9 January 2006 / Accepted: 19 January 2006 / Published online: 15 March 2006  
© Springer-Verlag 2006

**Abstract** Laser-induced breakdown spectroscopy (LIBS) has been applied to the analysis of three chromium-doped soils. Two chemometric techniques, principal components analysis (PCA) and neural networks analysis (NNA), were used to discriminate the soils on the basis of their LIBS spectra. An excellent rate of correct classification was achieved and a better ability of neural networks to cope with real-world, noisy spectra was demonstrated. Neural networks were then used for measuring chromium concentration in one of the soils. We performed a detailed optimization of the inputs of the network so as to improve its predictive performances and we studied the effect of the presence of matrix-specific information in the inputs examined. Finally the inputs of the network—the spectral intensities—were replaced by the line areas. This provided the best results with a prediction accuracy and precision of about 5% in the determination of chromium concentration and a significant reduction of the data, too.

**Keywords** Laser-induced breakdown spectroscopy (LIBS) · Soils · Neural networks analysis (NNA) · Principal components analysis (PCA) · Classification · Quantitative measurement

### Introduction

Laser-induced breakdown spectroscopy (LIBS) provides some significant advantages in the chemical analysis of

soils. The experimental setup is simple and compact, providing fast measurements in the tens-to-hundreds of ppm range without any preparation of the sample. These features make LIBS applicable for in situ or on-site analysis, which would be of great help in the chemical characterization of polluted soils. As LIBS is an atomic spectroscopy technique, it is particularly suited for the detection of metals, and several previous papers have demonstrated its applicability in this area [1–8].

In addition to this environmental context, the recent development of Echelle spectrometers [9–11] now permits the acquisition of broadband spectra at high resolution, i.e., extremely rich spectra comprising several thousands of points. This presents the challenge of finding an efficient way to extract the pertinent information from such data-rich spectra. This problem is complicated by the fact that soils are composite matrices containing multiple chemical elements at very different concentration levels. As a result, LIBS spectra of soils tend to be very complex and can exhibit strong spectral interferences.

Consequently a suitable advanced data treatment methodology is needed to extract both qualitative and quantitative information from LIBS spectra. In a previous study [12], we compared the performances of the standard calibration curve—i.e., the peak area versus the concentration—to two multivariate methods, partial least-squares regression (PLSR) and neural networks analysis (NNA), for the quantitative measurement of chromium concentration in soils. We clearly concluded that neural networks led to the best results in terms of prediction accuracy, prediction precision, and limit of detection, and that this approach to spectra analysis could really improve the analytical results obtained by LIBS.

This paper focuses on the use of neural networks for soils analysis. We first investigated the potential of networks to classify spectra of different soils and we compared their performances to the more standard principal components analysis (PCA). The question as to which part(s) of the spectra to use as inputs for neural networks for the optimal prediction of chromium concentration was then addressed.

Awarded a poster prize on the occasion of the Euro-Mediterranean Symposium on Laser-induced Breakdown Spectroscopy (EMSLIBS 2005), Aachen, Germany, 6–9 September 2005.

J.-B. Sirven (✉) · B. Bousquet · L. Canioni · L. Sarger ·  
S. Tellier · M. Potin-Gautier · I. Le Hecho  
Université Bordeaux 1,  
CPMOH, 351 Cours de la Libération,  
33405 Talence, France  
e-mail: jb.sirven@cpmoh.u-bordeaux1.fr  
Tel.: +33-540002785  
Fax: +33-540006970

**Table 1** Composition of Pierroton and Bourran soils

Constituent	Pierroton	Bourran
Sand (%)	94.19	28.18
Organic matter (%)	3.375	1.168
Ca (%)	0.113	0.362
Mg (%)	0.023	0.358
K (%)	0.325	1.843
Na (%)	0.091	0.61
Fe (%)	0.109	2.64
P (%)	0.011	0.051
Al (%)	0.593	0.544
Cd (ppm)	0.067	0.271
Cu (ppm)	3.51	22.21
Cr (ppm)	6.69	65.75
Hg (ppm)	0.025	0.059
Mn (ppm)	40.55	818.72
Ni (ppm)	2.79	26.59
Pb (ppm)	8.28	25.97
Zn (ppm)	10.35	62.88

## Experimental

### Experimental setup

A standard LIBS configuration was used in this study consisting of a Nd:YAG laser (Continuum Surelite) that delivers 7-ns pulses at 355 nm with an energy of 3 mJ and a repetition rate of 3 Hz. The relatively low repetition rate ensures that the interaction between the ejected particles and the following laser shot is negligible. Pulses are focused onto the sample through a 50-mm-focal-length lens. The focal spot has a diameter of approximately 50  $\mu\text{m}$ , and each laser shot ablates a few  $\mu\text{m}$  of material. A 38.1-mm-focal-length parabolic mirror associated with a 75-mm lens transmits the light emitted by the plasma into a bundle of 200- $\mu\text{m}$ -diameter optical fibers connected to the spectrometer (Oriel MS260i, 2,400 grooves/mm grating blazed at 400 nm). Its resolution power is about 2,500 at 360 nm, and it is equipped with an intensified camera (Andor Technology iStar). Detection is delayed by 150 ns with respect to the arrival of the laser pulse on the sample surface, and the gate width is 2  $\mu\text{s}$ . For analysis, a total of 100 laser pulses are accumulated for every spectrum as the sample is continuously moved in its plane during the

acquisition at a speed of about 500  $\mu\text{m s}^{-1}$ , ensuring no overlap between successive impacts.

### Samples and datasets

For the study, pressed pellets were prepared for three different soils:

- (i) Kaolinite ( $\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$ ): this sample is a “model soil”. It is a homogeneous white clay extracted from a deposit and has a grain size distribution between 1 and 10  $\mu\text{m}$ . Several oxides are present in the matrix with a concentration of < 1%, particularly  $\text{TiO}_2$ ,  $\text{Fe}_2\text{O}_3$ ,  $\text{FeO}$ ,  $\text{MnO}$ ,  $\text{MgO}$ ,  $\text{CaO}$ ,  $\text{Na}_2\text{O}$ ,  $\text{K}_2\text{O}$ ,  $\text{P}_2\text{O}_5$ , and  $\text{SO}_4$ .
- (ii) Bourran soil: this is a real agricultural soil with a very homogeneous, brown aspect.
- (iii) Pierroton soil: this is another real agricultural soil containing 94.2% of sand.

The composition of the two agricultural soils is given in Table 1. Their grain size distribution was intentionally not measured, since our ultimate goal is to perform on-site measurements.

Samples are doped with chromium at various concentrations by sorption in the liquid phase and continuous agitation for 24 h; they are then centrifuged and lyophilized, i.e., treated by a freeze-drying process. No binder is used. The pellets are then formed in a pelleting-press by applying 250  $\text{kg cm}^{-2}$  for 5 min.

Of the three different soils used in this study, the Bourran soil and the kaolinite give flat, texturally homogeneous pellets that produce smooth and very reproducible LIBS spectra (Fig. 2). By contrast, the Pierroton soil consists almost entirely of coarse-grained sand that does not readily generate satisfactorily aggregated pressed pellets, but rather pellets that are friable, spatially heterogeneous, and which produce LIBS spectra of highly variable emission line intensities.

The matrix of the spectra in lines at the different concentrations forms the dataset. When using neural networks, datasets are divided into two for each soil, one for the calibration of the network (calibration set) and the other for prediction of concentrations (validation set). Table 2 summarizes the composition of the different datasets.

The chromium lines of interest that were used for the quantitative measurements are at 357.87, 359.35, and 360.53 nm.

**Table 2** Datasets used for each soil

		Kaolinite	Bourran soil	Pierroton soil
Calibration set	Cr concentrations (ppm)	0, 10, 50, 100, 200, 300, 500	66, 76, 116, 166, 266, 566	7, 107, 207, 507
	Number of spectra per concentration	1	5	5
Validation set	Cr concentrations (ppm)	10, 100, 500	66, 76, 116, 166, 266, 566	7, 107, 207, 507
	Number of spectra per concentration	10	5	5

## Data processing

Two chemometrics techniques were used in this study, principal components analysis (PCA) and neural networks analysis (NNA). PCA [13] has been already applied to LIBS for identification/classification purposes [14–19] and involves representing the spectra in a new space with far fewer dimensions, keeping the information contained in the dataset. This dataset  $S$  (matrix of the spectra in lines) corresponds to the coordinates of the spectra in the initial base of the  $n$  wavelengths. The eigenvectors of the covariance matrix  $S'S$ , where  $S'$  is the transpose matrix, are the principal components (PC). They form a new base which has the following properties:

- (i) Each PC defines an axis of maximum variance of the whole dataset.
- (ii) The axes are orthogonal, i.e., the information contained in a PC is uncorrelated to the one contained in the other PCs.
- (iii) The eigenvalues associated with each PC characterize the amount of information that it contains.

Typically a few components, instead of the initial hundreds of wavelengths, will represent the quasi-totality of the spectral information. From the base defined by the PCs we can calculate the coordinates of the spectra in the new representation space. Spectra are then commonly projected onto the plane  $PC_2/PC_1$  containing the major part of the information.

For qualitative and quantitative analyzes we used a neural network based on the back-propagation algorithm [20]. Artificial neural networks are computer systems inspired by biological neurons. Processing is distributed over a large number of basic units called *perceptrons* or *nodes*. Owing to their original structure, neural networks exhibit the following specific properties which make them particularly powerful tools:

- (i) Parallel data processing.
- (ii) Learning /self-organization ability.
- (iii) Memory distribution over the units, which induces a strong tolerance to errors or even destruction of the units, and also to noise.
- (iv) Generalization ability.

The perceptron is described in Fig. 1. It consists of  $n$  inputs  $x_1, \dots, x_n$  converging in the unit. Each  $x_i$  is affected by a weight  $w_i$ , and the total input of the node is the weighted sum of the  $x_i$ . This number is compared to a threshold  $\theta$  and the perceptron applies a non-linear transfer function

(sigmoid) to the difference between the two. The neural network which we used is a three-layer perceptron, also represented in Fig. 1. In our case  $x_i$  values are the intensities at the  $n$  wavelengths, i.e., the input layer has  $n$  nodes. The hidden layer has a variable number of nodes, typically a few units. The output layer has only one node since we are interested in only one element, but multiple outputs are possible.

Weights are initially random. During the calibration phase each spectrum of the calibration set is successively injected into the network, then at each iteration the weights are dynamically changed as a function of the difference between the concentration calculated by the network and the true value. This procedure is repeated until the network reaches a good prediction ability over the concentration range. For prediction, the weights being fixed, an unknown spectrum is simply injected into the network which then calculates the concentration.

To assess the quantitative performances of the neural network, we calculate four figures:

- (i) Average relative error of calibration (%) to evaluate calibration accuracy.

$$REC(\%) = \frac{100}{N_C} \sum_{i=1}^{N_C} \left| \frac{\hat{c}_i - c_i}{c_i} \right|$$

where  $N_C$  is the number of calibration spectra,  $c_i$  is the true concentration, and  $\hat{c}_i$  is the estimated concentration.

- (ii) Average relative error of prediction (%) to evaluate prediction accuracy.

$$REP(\%) = \frac{100}{N_V} \sum_{i=1}^{N_V} \left| \frac{\hat{c}_i - c_i}{c_i} \right|$$

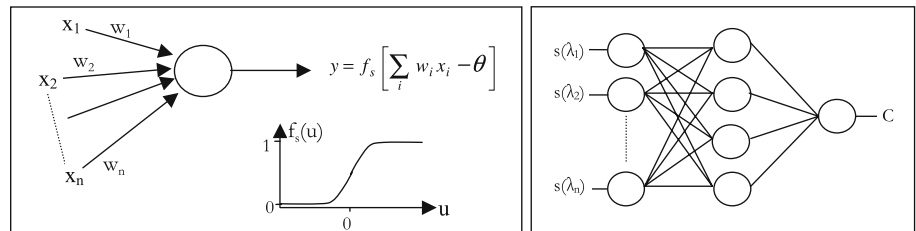
where  $N_V$  is the number of validation spectra.

- (iii) Average relative standard deviation of predicted concentrations of the validation set (%) to evaluate prediction precision.

$$RSD(\%) = \frac{100}{N_{conc}} \sum_{k=1}^{N_{conc}} \frac{\sigma_{C_k}}{c_k} \text{ with } \sigma_{C_k}^2 = \sum_{i=1}^p \frac{(\hat{c}_{ik} - c_k)^2}{p-1}$$

where  $N_{conc}$  is the number of different concentrations in the validation set,  $p$  is the number of spectra per concentration, and  $\sigma_{C_k}$  is the standard deviation obtained at concentration  $c_k$ .

**Fig. 1** Principle of the perceptron (*left*) and of a three-layer neural network (*right*).  $f_s$  is a sigmoid function



(iv) Limit of detection (ppm), calculated by fitting predicted concentrations of the validation set versus true concentrations by a straight line  $a+bc$ , so that the LOD is given by

$$LOD(ppm) = \frac{3\sigma_a}{b}$$

where  $\sigma_a$  is the standard deviation of  $a$  given by the regression.

## Results and discussion

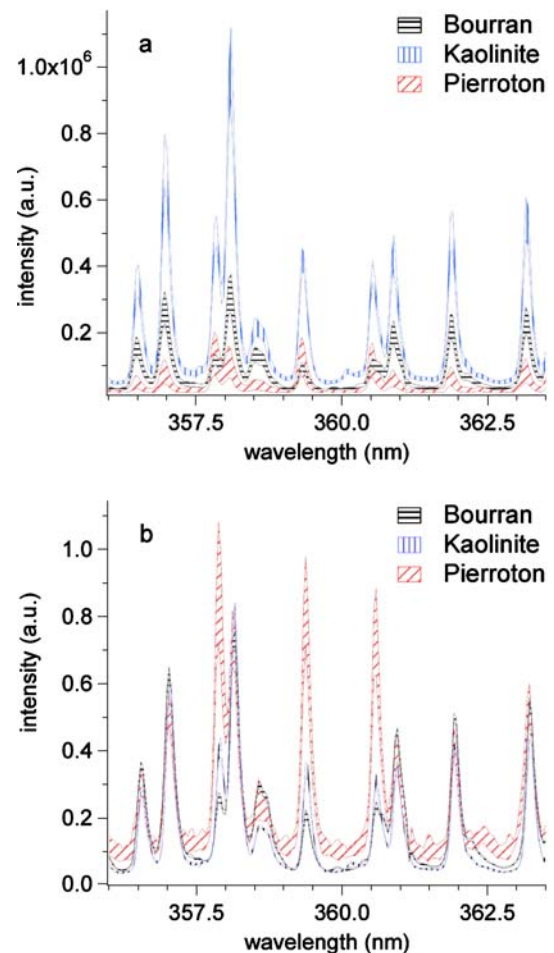
### Classification of soil spectra

Identification or classification of samples is often the first step of the analytical procedure after data acquisition. Indeed, if the prime objective is to measure an elemental concentration, quantitative techniques to process spectra are matrix-dependent. Even if multi-matrices analysis is possible [12], analytical performances will always be better when only one type of sample is measured, hence the interest in using techniques to classify samples prior to the determination of the concentration.

We have three different soils for this analysis. Since iron is present in the three matrices, spectra were normalized with respect to the 373.49-nm emission line of Fe. Figure 2 shows typical spectra of each soil.

We can see from Fig. 2a that raw spectra fluctuations are of the same order of magnitude for the three soils, although the relative variations of the Pierroton spectra seem more important. Figure 2b shows that normalizing by the Fe line compensates well the fluctuations of Bourran and kaolinite spectra, i.e., the *global* spectrum intensity changes from one spectrum to the other. In contrast, the normalization does not enable one to reduce Pierroton spectra variations. This means that *ratios* between lines change, i.e., matrix effects occur which cannot be counterbalanced by the normalization. However, thinking about a real-world application, it seemed interesting to us to check if we could identify these samples. If a quantitative measurement by LIBS is impossible with Pierroton spectra, at least we would be able to exclude them from the subsequent analysis.

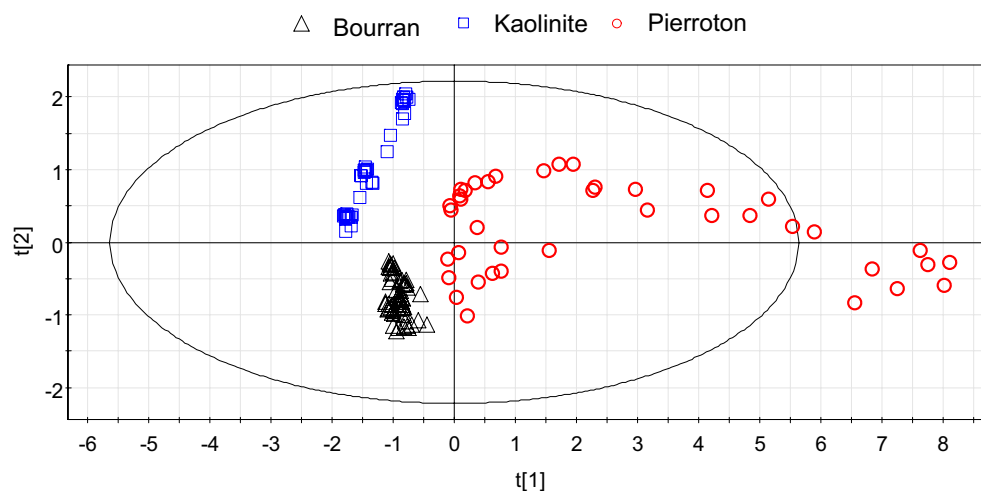
A 10-nm spectral window between 356 and 366 nm was selected to produce about 250 data points per spectrum. PCA was first employed to classify the samples. Figure 3 shows the coordinates of the spectra in the plane of the first two components  $t[1]$  and  $t[2]$  which contain 88.8% of the total spectral information. The ellipse in Fig. 3 corresponds to the Hotelling  $T^2$  with a confidence interval of 95%. It is an indication of how well the PC decomposition represents the data, i.e., spectra should lie inside this limit. As one can see, Bourran spectra form a well-defined cluster, whereas kaolinite spectra are aligned along a certain direction. This observation means that for this soil, both  $t[1]$  and  $t[2]$  are linked to chromium concentration. This is less clear for



**Fig. 2** Raw (a) and normalized (b) spectra of the three soils. Chromium concentration is 166 ppm for the Bourran soil, 100 ppm for the kaolinite, and 107 ppm for the Pierroton soil. The domain of variation on each side of the average spectrum is represented  $\pm$  one standard deviation over 10 spectra

Bourran samples, although one may detect a linear trend in the cluster indicating the same correlation, and it is clearly not the case for Pierroton samples. Yet when doing a PCA, we are not interested in chromium concentration but rather in separating the three soils. From this point of view, we can say from Fig. 3 that the three classes are correctly discriminated, but that there is a risk that new Pierroton spectra may be misclassified due to the important scattering of these spectra in the projection plane. In addition, a non-negligible fraction of these spectra lie outside the 95% confidence interval, which means that they are not well represented by the PCA decomposition. Furthermore, consideration of additional principal components did not provide a better discrimination of the Pierroton spectra. The influence of the spectral bandwidth on ability of PCA to classify the soils was also examined. No significant difference was observed between the entire 40-nm spectral window and a smaller window of 10 nm. Below this limit, input spectra contain insufficient information to distinguish the different matrices and the classification becomes less accurate. In conclusion, PCA enables one to sort spectra of

**Fig. 3** PCA of spectra from three different soils. The projection plane is the plane of the first two principal components



our three soils, but a doubt remains about its ability to correctly identify Pierroton spectra.

We also used neural networks to sort the same spectra, since the primary function of these networks is related to identification and classification problems [20]. For that purpose, arbitrary numerical values were assigned to each soil—0 for Kaolinite, 0.5 for Bourran, and 1 for Pierroton—but any values could be taken. The three-layer neural network had three nodes in the hidden layer and was first calibrated to predict these values from the calibration set of 57 spectra described in Table 2. The prediction was then performed with the validation set of 80 spectra. Results are presented in Table 3 for the validation set, which shows that predicted values are very close to target values, and that the standard deviation is less than 2% with respect to the initial range of values [0,1] in all three cases. With the exception of one Kaolinite sample, all predicted values lie within the target value  $\pm 0.05$ . If this tolerance is extended to  $\pm 0.1$ , all spectra are properly sorted by the network, leading to a 100% rate of correct classification. It is especially noticeable that Pierroton samples are perfectly separated from other soils. Therefore the conclusion is drawn that neural networks can be efficiently used in LIBS for classification purposes with an excellent sorting ability.

#### Measurement of Cr concentration

For the determination of chromium concentration in kaolinite samples by NNA, spectra were normalized with respect to the Fe line at 373.49 nm. The datasets used for calibration and validation were those described in Table 1.

**Table 3** Identification of three soils through neural networks prediction of arbitrary target values

Soil	Target value	Average predicted value	Standard deviation
Kaolinite	0	0.0136	0.0111
Bourran	0.5	0.5071	0.0145
Pierroton	1	0.9971	0.0012

In each case four specific parameters of the network were optimized during the calibration phase: number of hidden nodes, learning rate, momentum factor, and number of iterations. Given that there is a large quantity of information contained in the LIBS spectra, and that working with small networks is preferable as they have a better generalization ability and are less prone to overfitting [21] (i.e., having a prediction error larger than the calibration error), an important objective of this work was to determine the optimal input set for the network. Thus, to begin, the spectral range of the data was reduced in order to investigate its effect on calibration and prediction accuracy, prediction precision, and limit of detection (Table 4). Three intense chromium lines at 357.87, 359.35, and 360.53 nm were concatenated for the analysis shown in the last row of Table 4. In all other cases a spectral window centered at 360 nm was used.

Table 4 shows that the results are quite independent of the spectral bandwidth from 40 to 4 nm. No reduction of overfitting is observed either. The only significant discrepancy is the inexplicably high LOD associated with the 20-nm window. When using only the 359.35-nm line, which does not interfere with a matrix line in contrast to the two others, prediction is less accurate and less precise. This phenomenon is still amplified when the three Cr lines are concatenated: substantial overfitting is observed, although the network is rather small—calculations were done with two hidden nodes—and the LOD is high. This is possibly due to the fact that the two lateral chromium emission lines at 357.87 and 360.53 nm interfere with different matrix lines (Fe lines at 358.12 and 360.89 nm, Mn line at 360.85 nm). As a result their respective variations may be uncorrelated and cannot be compensated for by other parts of the spectra; hence, the calibration is more difficult than with only one line.

The results in Table 4 lead to the conclusion that dividing the number of inputs of the network by a factor of 10 does not have a large effect on the NNA prediction performances. However, it is necessary to have a certain amount of information in the LIBS spectra that is not correlated to chromium, since using only the Cr emission lines provides

**Table 4** Influence of spectral bandwidth on the analytical performances of the neural network

Spectral bandwidth	Number of points	REC (%)	REP (%)	RSD (%)	LOD (ppm)
40 nm (whole spectra)	1,024	2.4	5.4	6.8	28
20 nm	495	3.2	5.2	6.8	55
10 nm	248	2.6	5.5	7.4	29
4 nm	100	2.3	5.3	6.6	28
359.35 nm Cr line ( $\pm 0.5$ nm)	25	0.8	7.9	7.7	28
3 concatenated Cr lines	38	1.2	9.7	7.7	52

worse results in the prediction of the concentration. This means that the neural network performs its own signal normalization from regions in the spectra associated with the soil matrix. Due to self-absorption in the plasma, the relationship between the plasma emission and the concentration is non-linear, and the network needs information about the soil matrix in order to model this non-linearity.

Therefore, keeping in mind that the neural network should be as small as possible and that the number of inputs has to be optimized, the need to have matrix data in the inputs can be met simply by concatenating an additional spectral window containing matrix information to the three Cr lines, or to the single 359.35-nm line (Table 5). A 0.9-nm zone was then selected that includes the two Fe lines at 373.49 and 373.71 nm, not resolved by the spectrometer: this particular spectral window is henceforth referred as the matrix window. The effect of concatenating two matrix windows instead of one was also tested. The second window, centered at 344 nm, is 1.6-nm wide and includes the Fe emission lines triplet at 344.06, 344.10, and 344.39 nm, not resolved by the spectrometer either.

We see that considering one or more matrix windows (Table 5) and three Cr lines enables us to achieve equivalent results to those obtained with an unbroken spectral window (Table 4), with slightly better performances when using two additional windows instead of one. When using only one Cr emission line, results are not significantly improved compared to Table 4, except for the LOD. This means that the line signal at 359.35 nm after spectra normalization is quite reproducible and does not benefit from an additional normalization inside the network.

When tracing a calibration curve to measure a concentration by LIBS, or when using punctual detectors, the peak area is the quantity of interest. It is interesting to investigate

what happens with neural networks in this case. Regarding one emission line, we can consider that the peak area contains as much information as the spectral intensities, but it is concentrated on a single input node instead of several dozens, allowing reduction of the dimension of the network. The lower part of Table 5 presents calculations using the areas of the peaks, which were used as inputs for the neural network. Matrix information was also included in the network inputs by integrating the additional window around 373.5 nm over its spectral range, and only one input node was then added to the Cr peaks areas. Table 5 shows that the neural network performances are improved when using peaks areas instead of spectral intensities, and that overfitting is also reduced. Finally, the best results were obtained when entering the areas of the three Cr peaks and of one matrix-specific spectral window into the neural network, using four input nodes and 19 hidden ones. This resulted in a prediction accuracy of 4.2%, a prediction precision of 5.1%, and a LOD of 13 ppm. It is also worth mentioning that compared to the initial data—the whole spectra—a significant reduction of the number of input variables from 1,024 to 4 was achieved, together with a significant reduction of calculation time.

## Conclusion

The aim of this paper was to determine the chromium concentration of a soil using a neural networks analysis of LIBS spectra. As a first step in the analysis, the potential of neural networks for qualitative classification of three different soils was investigated and compared to PCA. Neural networks were shown to be more efficient, particularly in the case of real-world, noisy, and variable spectra, with a rate of correct identification of 100%.

**Table 5** Influence of concatenation of matrix windows to Cr lines (spectral intensities or peaks areas) on the analytical performances of the neural network

Network inputs	Number of points	REC (%)	REP (%)	RSD (%)	LOD (ppm)
3 Cr lines + 1 matrix window	60	3.0	5.7	6.2	27
3 Cr lines + 2 matrix windows	102	2.9	5.1	6.4	28
359.35 nm Cr line + 1 matrix window	48	2.8	7.4	8.0	19
3 Cr peaks areas + matrix window area	4	2.2	4.2	5.1	13
359.35 nm Cr peak area + matrix window area	2	2.5	5.9	7.2	19

Optimization of the input of the neural network was then undertaken to reduce its dimension, and therefore to reduce overfitting and to improve its performance for quantitative elemental analysis. The results obtained were quite independent of the spectral bandwidth and it was observed that soil matrix information had to be included in the neural network for optimal results. The use of matrix-specific spectral windows in addition to the chromium lines produced equivalent performances, with a reduced number of inputs. Finally, the lines spectra were replaced by their respective areas, which further improved the NNA results. A prediction accuracy and precision of 4–5% was obtained with a reasonably small neural network, and the number of input variables was reduced by more than two orders of magnitude. This work illustrates the power of neural networks as an advanced spectral treatment for the determination of elemental concentrations. We think that this innovative technique represents the future for competitive and mature quantitative LIBS.

**Acknowledgements** We gratefully thank Jean-Pierre Dubost (Université Bordeaux 2, France) for the fruitful discussions on chemometrics. This research is financially supported by ANR, ADEME and Conseil Régional d'Aquitaine.

---

## References

1. Wisbrun R, Schechter I, Niessner R, Schroeder H, Kompa KL (1994) *Anal Chem* 66:2964–2975
2. Lazic V, Barbini R, Colao F, Fantoni R, Palucci A (2001) *Spectrochim Acta Part B* 56:807–820
3. Mosier-Boss PA, Lieberman SH, Theriault GA (2002) *Environ Sci Technol* 36:3968–3976
4. Bustamante MF, Rinaldi CA, Ferrero JC (2002) *Spectrochim Acta Part B* 57:303–309
5. Hilbk-Kortenbruck F, Noll R, Wintjens P, Falk H, Becker C (2001) *Spectrochim Acta Part B* 56:933–945
6. Wainner RT, Harmon RS, Miziolek AW, McNesby KL, French PD (2001) *Spectrochim Acta Part B* 56:777–793
7. Capitelli F, Colao F, Provenzano MR, Fantoni R, Brunetti G, Senesi N (2002) *Geoderma* 106:45–62
8. Eppler AS, Cremers DA, Hickmott DD, Ferris MJ, Koskelo AC (1996) *Appl Spectrosc* 50:1175–1181
9. Bauer HE, Leis F, Niemax K (1998) *Spectrochim Acta Part B* 53:1815–1825
10. Detalle V, Héon R, Sabsabi M, St-Onge L (2001) *Spectrochim Acta Part B* 56:1011–1025
11. Fichet P, Menut D, Brennetot R, Vors E, Rivoallan A (2003) *Appl Opt* 42:6029–6035
12. Sirven JB, Bousquet B, Canioni L, Sarger L (2005) *Anal Chem* (accepted)
13. Adams MJ (2004) *Chemometrics in analytical spectroscopy*, 2nd edn. RSC, Cambridge, UK
14. Hybl JD, Lithgow GA, Buckley SG (2003) *Appl Spectrosc* 57:1207–1215
15. Goode SR, Morgan SL, Hoskins R, Oxsher A (2000) *J Anal At Spectrom* 9:1133–1138
16. Martin MZ, Labbé N, Rials TG, Wullschleger SD (2005) *Spectrochim Acta Part B* 60:1179–1185
17. Samek O, Telle HH, Beddows DCS (2001) *BMC Oral Health* 1:1
18. Samuels AC, De Lucia FC Jr, McNesby KL, Miziolek AW (2003) *Appl Opt* 42:6205–6209
19. Munson CA, De Lucia FC Jr, Piehler T, McNesby KL, Miziolek AW (2005) *Spectrochim Acta Part B* 60:1217–1224
20. Beale R, Jackson T (1991) *Neural computing: an introduction*. Adam Hilger, Bristol
21. Dieterle F, Busche S, Gauglitz G (2004) *Anal Bioanal Chem* 380:383–396